The problem is with the School of Computer Science, McGill University, 345 Sherbrooke

\[ \| Y \| = \text{sup} | f(y) - g(y) | \]

One of the classic criteria for simulation purposes? How should one measure the goodness of the approximation function F? One of the criteria is to have a good understanding of the consequences of the function F. Whenever the function F is used to simulate a complicated problem, the function F must be a good approximation of all the data it is used for. Sometimes one has difficulty finding a function F, but sometimes people (random variables) are available in less time or with less space. Instead, the function F is used as a function with distribution function G instead, the function F is used as a function with distribution function G. Consider the situation where one needs random variables with distribution function G. A Note on Approximations in

1. Introduction

Random Variate Generation

McGill University
LUC DEBOER
covered for the generation of random variables with density \( f \). Similarly, consider the case of the distribution of \( x \) if the distribution method is conditional on the values of the random variables. Then, if \( x \) is close to \( f \), the conditional density \( f_{|x} \) would be a delta function. The conditional density is then given by \( f_{|x} = \delta(x) \).

### 2. Practical Inequalities

In this section, we introduce and discuss a technique for solving inequalities and in particular inequalities in the determination of upper bounds on the solution space. This technique is usually referred to as the compact method, as it is very simple to compute. In this section, we give several examples of the compact method, which can be demonstrated without much effort. We\'ll discuss the following inequalities:

- For all \( x \), \( y \):
  \[ |x - y| \leq |x + y| \]  
- For all \( x \), \( y \), \( z \):
  \[ |x - y| + |y - z| + |z - x| = 2 \]

\( \text{Let } f \text{ and } g \text{ be densities on } \mathbb{R}, \text{ and let us define the following constants:} \)

\[ \mathcal{A}(f, g) = \int_{\mathbb{R}} f(x) g(x) \, dx \]

\[ \mathcal{B}(f, g) = \int_{\mathbb{R}} f(x) \, dx \cdot \int_{\mathbb{R}} g(x) \, dx \]

\[ \mathcal{C}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{D}(f, g) = \int_{\mathbb{R}} f(x) \, dx \cdot \int_{\mathbb{R}} g(x) \, dx \]

\[ \mathcal{E}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{F}(f, g) = \int_{\mathbb{R}} |f(x) - g(x)| \, dx \]

\[ \mathcal{G}(f, g) = \int_{\mathbb{R}} |f(x) + g(x)| \, dx \]

\[ \mathcal{H}(f, g) = \int_{\mathbb{R}} f(x)^2 \, dx \cdot \int_{\mathbb{R}} g(x)^2 \, dx \]

\[ \mathcal{I}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{J}(f, g) = \int_{\mathbb{R}} |f(x) + g(x)| \, dx \]

\[ \mathcal{K}(f, g) = \int_{\mathbb{R}} |f(x) - g(x)| \, dx \]

\[ \mathcal{L}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{M}(f, g) = \int_{\mathbb{R}} |f(x) + g(x)| \, dx \]

\[ \mathcal{N}(f, g) = \int_{\mathbb{R}} |f(x) - g(x)| \, dx \]

\[ \mathcal{O}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{P}(f, g) = \int_{\mathbb{R}} |f(x) - g(x)| \, dx \]

\[ \mathcal{Q}(f, g) = \int_{\mathbb{R}} |f(x) + g(x)| \, dx \]

\[ \mathcal{R}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{S}(f, g) = \int_{\mathbb{R}} |f(x) + g(x)| \, dx \]

\[ \mathcal{T}(f, g) = \int_{\mathbb{R}} |f(x) - g(x)| \, dx \]

\[ \mathcal{U}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{V}(f, g) = \int_{\mathbb{R}} |f(x) - g(x)| \, dx \]

\[ \mathcal{W}(f, g) = \int_{\mathbb{R}} |f(x) + g(x)| \, dx \]

\[ \mathcal{X}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{Y}(f, g) = \int_{\mathbb{R}} |f(x) + g(x)| \, dx \]

\[ \mathcal{Z}(f, g) = \int_{\mathbb{R}} |f(x) - g(x)| \, dx \]

\[ \mathcal{A}(f, g) = \int_{\mathbb{R}} f(x) g(x) \, dx \]

\[ \mathcal{B}(f, g) = \int_{\mathbb{R}} f(x) \, dx \cdot \int_{\mathbb{R}} g(x) \, dx \]

\[ \mathcal{C}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{D}(f, g) = \int_{\mathbb{R}} f(x) \, dx \cdot \int_{\mathbb{R}} g(x) \, dx \]

\[ \mathcal{E}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{F}(f, g) = \int_{\mathbb{R}} |f(x) - g(x)| \, dx \]

\[ \mathcal{G}(f, g) = \int_{\mathbb{R}} |f(x) + g(x)| \, dx \]

\[ \mathcal{H}(f, g) = \int_{\mathbb{R}} f(x)^2 \, dx \cdot \int_{\mathbb{R}} g(x)^2 \, dx \]

\[ \mathcal{I}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{J}(f, g) = \int_{\mathbb{R}} |f(x) + g(x)| \, dx \]

\[ \mathcal{K}(f, g) = \int_{\mathbb{R}} |f(x) - g(x)| \, dx \]

\[ \mathcal{L}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{M}(f, g) = \int_{\mathbb{R}} |f(x) + g(x)| \, dx \]

\[ \mathcal{N}(f, g) = \int_{\mathbb{R}} |f(x) - g(x)| \, dx \]

\[ \mathcal{O}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{P}(f, g) = \int_{\mathbb{R}} |f(x) - g(x)| \, dx \]

\[ \mathcal{Q}(f, g) = \int_{\mathbb{R}} |f(x) + g(x)| \, dx \]

\[ \mathcal{R}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{S}(f, g) = \int_{\mathbb{R}} |f(x) + g(x)| \, dx \]

\[ \mathcal{T}(f, g) = \int_{\mathbb{R}} |f(x) - g(x)| \, dx \]

\[ \mathcal{U}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{V}(f, g) = \int_{\mathbb{R}} |f(x) - g(x)| \, dx \]

\[ \mathcal{W}(f, g) = \int_{\mathbb{R}} |f(x) + g(x)| \, dx \]

\[ \mathcal{X}(f, g) = \left( \int_{\mathbb{R}} f(x) \, dx \right)^2 + \left( \int_{\mathbb{R}} g(x) \, dx \right)^2 \]

\[ \mathcal{Y}(f, g) = \int_{\mathbb{R}} |f(x) + g(x)| \, dx \]

\[ \mathcal{Z}(f, g) = \int_{\mathbb{R}} |f(x) - g(x)| \, dx \]
Inequality + Let Δ and G have densities and Θ, \( \Theta \), on \( \mathbb{R} \). Then

\[
\left[ 1 - x^p \left( \frac{(x)^{\gamma}}{\|x\|^\beta} \right) \right] \leq \frac{\Delta}{\Gamma}
\]

Finally, since an \( \varepsilon \) norm is often easier to handle than an \( \| x \| \) norm, we

\[
x^p \left( \frac{(x)^{\gamma}}{\|x\|^\beta} \right) \leq x^p \left( \frac{(x)^{\gamma}}{\|x\|^\beta} - 1 \right) \leq \frac{\Delta}{\Gamma}
\]

It should be noted that since \( -1 \leq x \leq 1 \), we always have

\[
x^p \left( \frac{(x)^{\gamma}}{\|x\|^\beta} \right) \leq \frac{\Delta}{\Gamma}
\]

Inequality (Kullback, 1959), Let \( \Delta \) and \( G \) have densities and \( \Theta \), \( \Theta \), on \( \mathbb{R} \). Then

\[
\left[ 1 - x^p \left( \frac{(x)^{\gamma}}{\|x\|^\beta} \right) \right] \leq \frac{\Delta}{\Gamma}
\]

Another inequality that is often easily compared is

\[
\left[ 1 - x^p \left( \frac{(x)^{\gamma}}{\|x\|^\beta} \right) \right] \leq \frac{\Delta}{\Gamma}
\]

Proof. Notice that

\[
\left[ 1 - x^p \left( \frac{(x)^{\gamma}}{\|x\|^\beta} \right) \right] \leq \frac{\Delta}{\Gamma}
\]

Inequality (Kullback, 1959), Let \( \Delta \) and \( G \) have densities and \( \Theta \), \( \Theta \), on \( \mathbb{R} \). Then

\[
\left[ 1 - x^p \left( \frac{(x)^{\gamma}}{\|x\|^\beta} \right) \right] \leq \frac{\Delta}{\Gamma}
\]

Another inequality that is often easily compared is

\[
\left[ 1 - x^p \left( \frac{(x)^{\gamma}}{\|x\|^\beta} \right) \right] \leq \frac{\Delta}{\Gamma}
\]

Proof. Notice that

\[
\left[ 1 - x^p \left( \frac{(x)^{\gamma}}{\|x\|^\beta} \right) \right] \leq \frac{\Delta}{\Gamma}
\]

Inequality (Kullback, 1959), Let \( \Delta \) and \( G \) have densities and \( \Theta \), \( \Theta \), on \( \mathbb{R} \). Then

\[
\left[ 1 - x^p \left( \frac{(x)^{\gamma}}{\|x\|^\beta} \right) \right] \leq \frac{\Delta}{\Gamma}
\]
4. THE NORMAL APPROXIMATION

which would minimize \( \mathbb{E} \) in most cases, see the previous remark

\[
\left(1 - \frac{d}{1 - x} - 1\right) \max
\]

cases, we may try to minimize

when \( \mathbb{E} \) is the upper bound for \( x \); see inequality (1); in some

\[
\left(1 - \frac{d}{1 - x} - 1\right) \max \mathbb{E}
\]

we may try to minimize

parameters in view of the fact that \( d \) and \( y \) are simple functions of only

cannot be improved upon except possibly in some particular cases

\[
\left(1 - \frac{d}{1 - x} - 1\right) \max \mathbb{E}
\]

Clearly,

\[
\frac{(x) f(x+y)}{(x) f(x)} \frac{(x) f(x+y)}{(x) f(x)} = \frac{(x) f(x+y)}{(x) f(x)}
\]

Define further that

\[
\frac{(x) f(x+y)}{(x) f(x)} \frac{(x) f(x+y)}{(x) f(x)} = \frac{(x) f(x+y)}{(x) f(x)}
\]

which is known as the Lebesgue density theorem; see Theorem and

\[
\frac{(x) f(x+y)}{(x) f(x)} \frac{(x) f(x+y)}{(x) f(x)} = \frac{(x) f(x+y)}{(x) f(x)}
\]

we may try to minimize

when \( f \) is both continuous and strictly positive on \( x \), then

\[
\left(1 - \frac{d}{1 - x} - 1\right) \max \mathbb{E}
\]

is valid. Define further that

\[
\frac{(x) f(x+y)}{(x) f(x)} \frac{(x) f(x+y)}{(x) f(x)} = \frac{(x) f(x+y)}{(x) f(x)}
\]

\[
\frac{(x) f(x+y)}{(x) f(x)} \frac{(x) f(x+y)}{(x) f(x)} = \frac{(x) f(x+y)}{(x) f(x)}
\]

\[
\frac{(x) f(x+y)}{(x) f(x)} \frac{(x) f(x+y)}{(x) f(x)} = \frac{(x) f(x+y)}{(x) f(x)}
\]

is valid. Define further that

\[
\frac{(x) f(x+y)}{(x) f(x)} \frac{(x) f(x+y)}{(x) f(x)} = \frac{(x) f(x+y)}{(x) f(x)}
\]

we may try to minimize

when \( f \) is both continuous and strictly positive on \( x \), then

\[
\left(1 - \frac{d}{1 - x} - 1\right) \max \mathbb{E}
\]

for \( \lambda = 0 \).

For "suitable" \( \lambda \), we have that \( \lambda = 0 \) for such densities.

3. A RELATIVE ERROR CRITERION

\[
\left(1 - \frac{d}{1 - x} - 1\right) \max \mathbb{E}
\]

we may try to minimize

when \( f \) is both continuous and strictly positive on \( x \), then

\[
\left(1 - \frac{d}{1 - x} - 1\right) \max \mathbb{E}
\]

for \( \lambda = 0 \).

For "suitable" \( \lambda \), we have that \( \lambda = 0 \) for such densities.

Next, by taking care that the supports do not contain \( x \) with positive

155
\[
\frac{\lambda_0}{\lambda} = \frac{1}{\sum_{n=1}^{\infty} \lambda^n} \leq 1
\]

then

X is approximated by a Poisson (\(\lambda_0/\lambda\)) distribution. For example, when \(\lambda = \frac{1}{2}\), we have:

\[
\frac{1}{\sum_{n=1}^{\infty} \lambda^n} = \frac{1}{\lambda} = \frac{1}{\frac{1}{2}} = 2
\]

where \(X\) is the sum of \(n\) independent Poisson random variables with \(X_i\) being the number of events in the interval \([0, t_i)\) for \(i = 1, 2, \ldots, n\).

6. THE POISSON APPROXIMATION

The Poisson distribution is a discrete probability distribution that models the number of events occurring in a fixed interval of time or space. It is commonly used to model the number of times a rare event occurs in a given time period or in a specific area. The Poisson distribution is characterized by a single parameter, \(\lambda\), which is the average number of events in the interval.

The Poisson distribution is often used as an approximation to the binomial distribution when the number of trials is large and the probability of success is small. This approximation is known as the Poisson approximation to the binomial distribution.

\[
P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}
\]

where \(X\) is the number of events, \(k\) is the number of successes, \(\lambda\) is the average number of events, and \(e\) is the base of the natural logarithm.

The Poisson distribution is often used in applications such as insurance, quality control, and biology to model the number of events that occur in a given interval. It is also used in traffic engineering to model the number of vehicles that pass a certain point in a given time period.

The Poisson distribution is characterized by a single parameter, \(\lambda\), which is the average number of events in the interval. The distribution is often used to model the number of events that occur in a given time period or in a specific area. The Poisson distribution is often used as an approximation to the binomial distribution when the number of trials is large and the probability of success is small.